

**REMARKS**

Claims 1-7, 9-21, 23-44, 46, 47, 49-65, 67-76, 78-92, 95-100, 102, 103, 105-112, 114-116 are currently pending in the subject application and are presently under consideration. The below comments present in greater detail distinctive features of applicants' claimed invention over the cited art that were conveyed to the Examiner over the telephone on February 5, 2008.

Favorable reconsideration of the subject patent application is respectfully requested in view of the comments herein.

**I. Rejection of Claims 1-7, 9-21, 23-44, 46, 47, 49-65, 67-76, 78-92, 95-100, 102, 103, 105-112 and 114-116 Under 35 U.S.C. §103(a)**

Claims 1-7, 9-21, 23-44, 46, 47, 49-65, 67-76, 78-92, 95-100, 102, 103, 105-112 and 114-116 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Bailey *et al.* (U.S. 20060167864) in view of Albion *et al.* (U.S. 20040240388). Withdrawal of this rejection is requested for at least the following reasons. Bailey *et al.* and Albion *et al.* alone or in combination fail to teach or suggest all features set forth in the subject claims.

Applicants' claimed subject matter relates to data analysis, and systems and methods for obtaining information from a networked system utilizing a distributed web crawler. Information gathered by a server's web crawler is compared to data retrieved by clients of the server to update the crawler's data. In particular, independent claim 1 recites *a data analysis system, comprising: a first component associated with a server of the data analysis system that facilitates generation of a first data set related to web page information obtained via a communication system; and a second component that coordinates a second data set relating to web page information from at least one distributed resource associated with at least a client of the server which interacts with the communication system.* Independent claims 37 further recites *refining the second data set to reflect information obtained from the third data set by adding unknown information to the second data set when new information is received from the distributed source via the third data set.* Independent claims 57, 61, 92 and 113 recite similar features. Bailey *et al.* and Albion *et al.* are silent regarding such novel features.

Bailey *et al.* relates to a search engine system for locating web pages with product offerings. At page 3 of the Final Office Action, the Examiner contends that Bailey *et al.* discloses such novel features of applicants' claimed invention. Applicants' representative avers

to the contrary. In accordance with the subject invention, a server hosts a web crawler that searches a communication network such as the Internet for other servers hosting web pages, gathers information about these web pages and compiles them for utilizing with a web page search engine (*See* applicants' Fig.1 and Fig.2). The server then sends a representation of this web page information to a client of the server. When the client accesses that particular web page or detects web pages that are unknown to the server, the client compiles changes/status and/or new information about the known and unknown web pages. This information is then transmitted to the server, which utilizes the information to update its original crawler web page data to reflect a new web page or change of contents in a known web page. At the cited portions, Bailey *et al.* discloses a web server application that processes user requests to query and make purchases from a catalog, *via* the internet 120. The web server records the user transactions within a query log. Further, Bailey *et al.* discloses the Product Spider database that has product scores and category ranking information about independent web sites unaffiliated with the host web site, that offer products for sale. When updating the database, URL's of the existing database are submitted to the second crawling stage, updated, duplicate submissions are detected and removed. However, the cited document is silent regarding utilizing web page information communicated by a client of the distributed web crawler system to update its original crawler web page data to reflect a new web page or change of contents in a known web page. For example, Bailey *et al.* teaches a conventional web crawler implemented by a server but does not teach or suggest that the web crawler 160 is updated with inputs from the clients 110 (*See* Bailey *et al.* Fig. 1 and paragraph [0037]) Thus Bailey *et al.* does not disclose a distributed web crawler wherein a client updates web pages associated with a server of the distributed system as recited by the subject claims.

Albion *et al.* relates to dynamic assignment of timers in a network transport engine that provides a connection between two applications running on different system interconnected via the network. Timer logic includes a counter, a crawler, a memory and a list of available timers. The crawler processes client request and keeps note of the timer information in the timer list located in the memory. Upon a request from the client, timers are allocated, de-allocated and restarted. Accordingly the timer list is updated. However, the second data set from the client is not web page information communicated by a client of the distributed web crawler system as recited by the subject claims. Thus, crawler 204 of Fig. 2 in Albion is a component that manages

timers accessed by clients rather than a component that provides unknown information or updates information when changes have occurred in the contents of the web page information as recited in the subject independent claims (*See e.g.*, Albion paragraph [0018]). Therefore, it is concluded that Albion *et al.* is silent regarding *refining the second data set to reflect information obtained from the third data set by adding unknown information to the second data set when new information is received from the distributed source via the third data set* as recited by the subject claims.

By distributing the web crawler functionality among the search server and its clients, the server utilizes the clients to obtain information from web page servers to facilitate in refining its own information. This helps in providing a more up-to-date, robust and spoof-proof data set from which a search engine can utilize data.

In view of at least the foregoing, it is readily apparent that both Bailey *et al.* and Albion *et al.* fail to teach or suggest all limitations of the claimed invention. Accordingly, it is respectfully requested that rejection of independent claims 1, 37, 57, 61, 92 and 113 (and the claims that depend there from) be withdrawn.

**CONCLUSION**

The present application is believed to be in condition for allowance in view of the above comments. A prompt action to such end is earnestly solicited.

In the event any fees are due in connection with this document, the Commissioner is authorized to charge those fees to Deposit Account No. 50-1063 [MSFTP475US].

Should the Examiner believe a telephone interview would be helpful to expedite favorable prosecution, the Examiner is invited to contact applicants' undersigned representative at the telephone number below.

Respectfully submitted,

AMIN, TUROCY & CALVIN, LLP

/Himanshu S. Amin/

Himanshu S. Amin

Reg. No. 40,894

AMIN, TUROCY & CALVIN, LLP  
24<sup>TH</sup> Floor, National City Center  
1900 E. 9<sup>TH</sup> Street  
Cleveland, Ohio 44114  
Telephone (216) 696-8730  
Facsimile (216) 696-8731